

基于自适应特征增强和生成器路径交互的 红外与可见光图像融合

杨叶君¹, 刘刚^{1*}, 肖刚², 顾新杰¹

¹上海电力大学自动化工程学院, 上海 200090;

²上海交通大学航空航天学院, 上海 200240

摘要 在基于生成对抗网络的红外与可见光图像融合中, 针对融合结果中不同物体之间边缘模糊、源图像信息提取不足和融合信息不平衡等问题, 提出一种基于自适应特征增强和生成器路径交互的深度融合方法。首先, 自适应增强块根据权重图锐化源图像中不同物体的边缘信息, 在自适应调整的特征增强损失与强度损失及梯度损失联合约束后, 融合图像的对比度与纹理细节可以同时得到增强。其次, 生成器路径交互结构在两条主路径之间增加交互卷积层以充分提取源图像信息, 采用密集连接的卷积网络可以加强特征图的传输。此外, 以主辅思想设计的内容损失函数和网络结构中引入的双鉴别器保证了融合结果中对比度与纹理细节的平衡。实验结果表明, 所提方法的融合结果无论是在主观视觉评价还是客观定量评价方面都非常有竞争力, 并且与其他方法相比, 所提方法的运行速度最快。

关键词 图像处理; 图像增强; 生成对抗网络; 自适应机制; 路径交互

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP222204

Infrared and Visible Image Fusion Based on Adaptive Feature Enhancement and Generator Path Interaction

Yang Yejun¹, Liu Gang^{1*}, Xiao Gang², Gu Xinjie¹

¹School of Automation Engineering, Shanghai University of Electric Power, Shanghai 200090, China;

²School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract For infrared and visible image fusion based on generation countermeasure networks, a deep fusion method based on adaptive feature enhancement and generator path interaction is proposed to solve problems, such as edge blurring between different objects in the fusion result, insufficient extraction of source image information, and imbalance of fusion information. First, the adaptive enhancement block sharpens the edge information of different objects in the source image according to the weight map. After the adaptive feature enhancement loss, intensity loss, and gradient loss are jointly constrained, the contrast and texture details of the fused image can be enhanced simultaneously. Second, the generator path interaction structure can fully extract the source image information by adding an interactive convolution layer between the two main paths, and the transmission of the feature map can be enhanced using a densely connected convolution network. In addition, a content loss function, designed based on the primary, secondary, and dual discriminator introduced in the network structure, ensures the balance between contrast and texture details in the fusion results. Experimental results show that the proposed method has very competitive results in both subjective visual and objective quantitative evaluations and is faster than the other methods.

Key words image processing; image enhancement; generative adversarial network; adaptive mechanism; path interaction

1 引言

图像融合的目的是将不同传感器捕获的图像信息

进行整合, 从而产生更高质量的融合图像。其中, 红外与可见光图像融合的应用非常广泛^[1]。红外图像由红外传感器捕获, 记录不同物体发出的热辐射, 被广泛应用

收稿日期: 2022-08-02; 修回日期: 2022-09-01; 录用日期: 2022-09-27; 网络首发日期: 2022-10-07

基金项目: 国家自然科学基金(61673270, 61203224)、国家重点基础研究发展规划项目 973 计划(2014CB744903)、上海市地方能力建设项目(2201050300)、上海浦江人才计划(16PJD028)

通信作者: *liugang@shiep.edu.cn

于目标检测与定位。可见光图像由可见光传感器捕获反射光而产生,其特点在于纹理细节丰富,符合人眼视觉感知。因此,采用红外与可见光图像融合方法对图像特征进行处理后可以生成具有显著对比度和丰富纹理细节的融合结果。此外,基于图像特征融合的方法也在目标检测和线结构光成像等领域得到了广泛应用^[2-4]。

根据图像融合方法的实现原理,可将其分为基于信号处理的传统方法和基于深度学习的方法^[5]。传统方法通常使用相关的数学变换提取源图像的特征,根据手动设计的融合策略融合特征,对融合后的特征进行反变换得到最终的结果。传统方法包括基于多尺度变换的方法^[6-7]、基于稀疏表示的方法^[8-9]、基于显著性的方法^[10]和混合方法^[11-12]等。然而,传统方法通过固定的数学变换提取的特征通常无法全面表达,没有考虑源图像之间的模态差异,限制了融合的性能。基于深度学习的方法依靠神经网络强大的非线性拟合与特征提取能力,在很大程度上弥补了传统方法的不足。根据网络架构,基于深度学习的方法可以分为基于卷积神经网络(CNN)、基于自动编码器(AE)和基于生成对抗网络(GAN)的方法。近年来,文献[13]和文献[14]相继提出基于CNN的统一图像融合框架以实现各种图像融合任务。但是针对红外与可见光图像融合:前者的融合结果偏向于可见光图像,红外热辐射信息明显不足;后者的融合结果偏向于保留热辐射信息,缺少足够的纹理细节。在基于AE的方法中,文献[15]通过一种改进的无参考图像评价指标计算损失函数以实现网络的无监督训练,而文献[16]和文献[17]使用预先训练好的AE网络来提取特征和重建图像,但是以上三者均使用传统的融合策略融合特征。在基于GAN的融合方法中,文献[18]提出一种端到端的融合模型,避免了像传统方法那样手动设计复杂的活动水平测量和融合规则,融合结果偏向于锐化的红外图像。文献[19]针对多分辨率图像融合,在融合网络

中应用了双鉴别器的条件GAN结构,迫使融合图像同时保留红外图像的热目标和可见光图像的纹理细节。文献[20]针对不同类型的源图像所包含的主次信息,设计了特定的内容损失函数,并且使用了带有多分类器的GAN模型以解决不平衡融合问题。然而,在上述基于GAN的方法所得到的融合图像中,纹理细节和热目标的边缘相对模糊、难以界定。

针对上述问题,本文提出一种基于自适应特征增强和生成器路径交互的红外与可见光图像融合方法。首先,在以往基于GAN的方法中,融合图像的纹理细节和热目标的边缘模糊,为此,自适应增强块通过权重图与强度损失联合约束不仅能锐化增强融合图像的纹理细节和热目标的边缘,还能进一步突出来自源图像的对比度和梯度信息。其次,生成器的网络得到了优化,生成器路径交互结构通过在两条主路径之间交换卷积信息以提取更多的源图像信息,同时密集连接的卷积网络加强了特征图的传输,使融合图像的质量更高。引入的双鉴别器使融合图像同时保留两种源图像中最重要的特征信息,保证了融合结果中各类信息的平衡。此外,自适应增强块通过权重图作用后的损失函数可以灵活调整。实验结果表明,由所提方法得到的融合图像边缘纹理清晰、热目标突出、表现更自然。

2 融合方法

2.1 网络总体框架

图像融合研究的重点在于如何保留源图像中显著的对比度与丰富的纹理细节,然而神经网络在图像生成中常见的平滑效应^[21]会导致融合结果的纹理细节不足。基于上述考虑,在融合图像的生成过程中,有必要引入调整损失函数的自适应机制来引导图像生成,为此而设计的生成对抗网络能增强融合图像的对比度和纹理细节,从而突出来自源图像的纹理信息和热目标的边缘。所提网络的整体框架如图1所示,将生成器

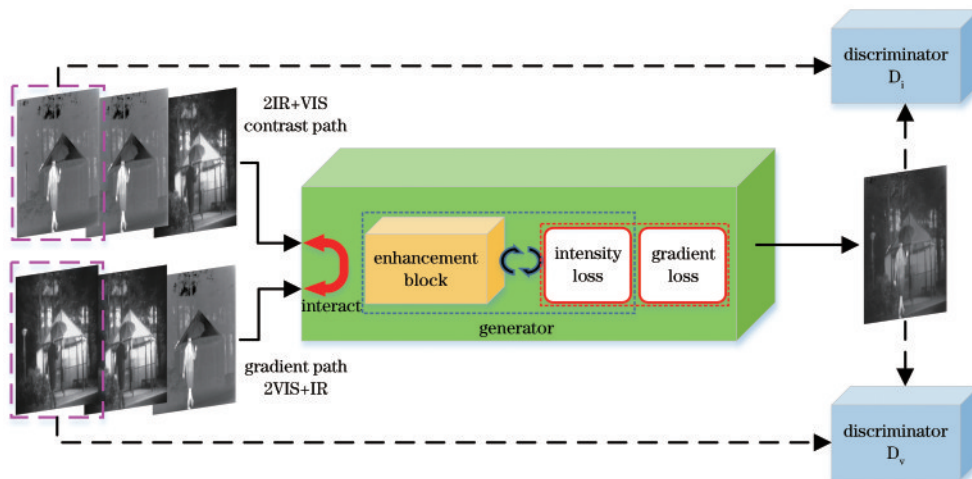


图1 基于自适应特征增强和生成器路径交互的融合网络框架

Fig. 1 Fusion network architecture based on adaptive feature enhancement and generator path interaction

的输入分为对比度路径与梯度路径,并在两条路径之间进行信息交互。在任一条路径中,以差值比连接形式构建的输入可以迫使网络不均等地提取对比度或者梯度信息^[20],从而保证在此路径下充分提取来自两种源图像的同目标信息。生成器由特定梯度损失、特定强度损失和自适应增强块联合约束,随即产生一个真实的融合图像以骗过两个鉴别器,两个鉴别器又不断地对融合图像与源图像进行区分,连续的对抗性学习引导生成器产生对比度高且纹理细节丰富的融合图像。

2.2 自适应增强块

所设计的自适应增强块采用不同半径组合的高斯卷积核对源图像的梯度图进行滤波,得到相应的权重图。自适应增强块根据权重图锐化来自源图像的纹理细节与热目标的边缘,突出融合图像的对比度和

纹理细节特征,如图 2 所示。在权重图中,针对源图像各类物体的边缘区域权重较大,这帮助网络在训练过程中更加关注各类物体的边缘。权重图作用于构建的特定强度损失函数,帮助优化融合图像的特征。换句话说,增强块可以自适应地引导融合过程锐化增强红外图像与可见光图像的目标物体。具体地:在权重图中,接近两幅源图像中各类物体边缘的区域,权重值逐渐变大;远离各类物体边缘的区域,权重值逐渐减小直至为零,这帮助边缘区域两侧的权重值得到平滑过渡。在增强块通过权重图结合强度损失后,生成器可以尽可能地保留并增强红外与可见光图像各自的热目标和纹理细节。在以往的深度学习中,损失函数是固定不变的,而在自适应增强块的作用下,损失函数得到了灵活调整,帮助不同的源图像各自的目标物体得到锐化与增强。

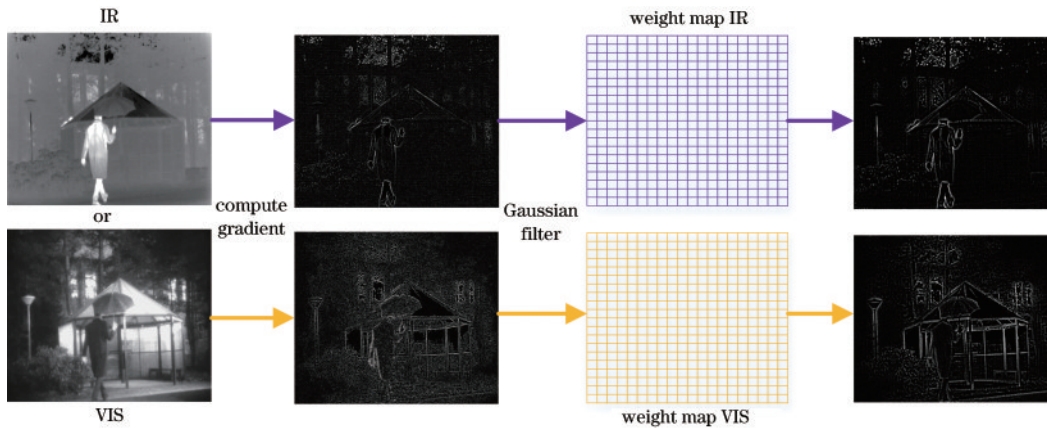


图 2 自适应增强块的原理图

Fig. 2 Schematic of adaptive enhancement block

2.3 生成器网络框架

生成器网络在输入端分为两条路径,分别提取红

外图像与可见光图像各自的特征,它的内部结构如图 3 所示。由于红外图像主要包含对比度信息,也有

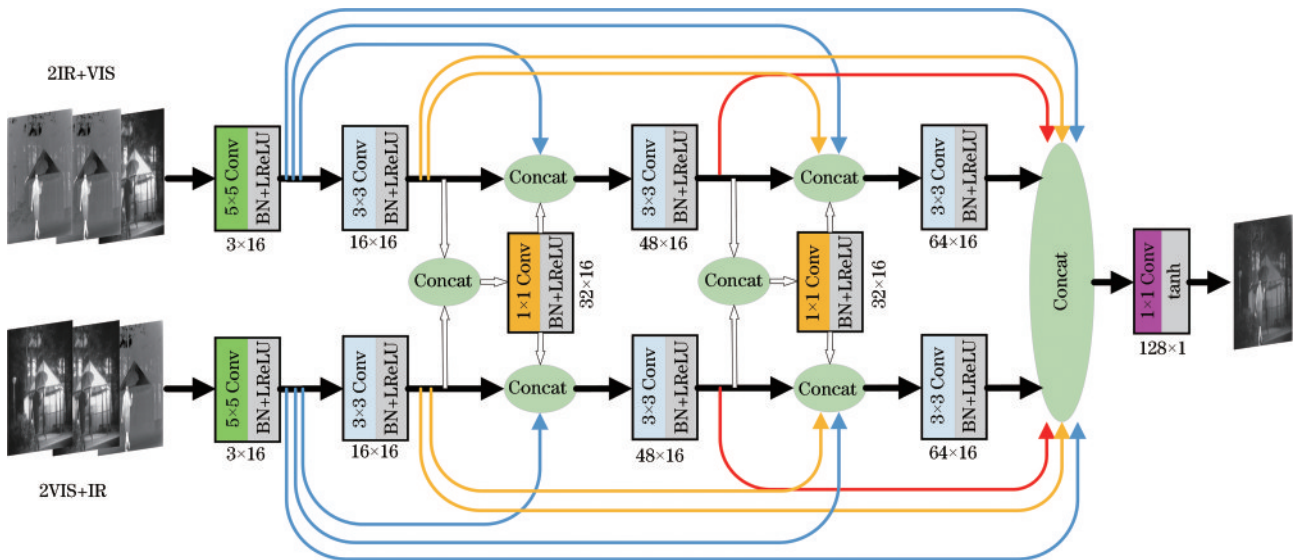


图 3 生成器网络框架

Fig. 3 Network architecture of generator

部分纹理细节信息,而可见光图像主要包含纹理细节信息,其次也有部分对比度信息。因此,以差值比连接的形式来构建输入。对于对比度路径,以沿通道连接的 2 幅红外图像与 1 幅可见光图像作为输入。同理,对于梯度路径,以沿通道连接的 2 幅可见光图像和 1 幅红外图像作为输入。

对于两条路径中任一条路径,前 4 个卷积层用来提取特征,其中,第 1 个卷积层使用 5×5 的卷积核,其余 3 个卷积层均使用 3×3 的卷积核,每个卷积层通过各自的过滤器获得 16 个特征图。为了弥补特征损失并且重用之前卷积层计算的特征,网络中两条路径均采用 DenseNet^[22],即每 1 个卷积层的输入都是来自这一条路径中所有之前卷积层的输出。此外,为了提取更多的源图像信息,在两条路径之间进行信息交互,信息交互的卷积核大小是 1×1 ,交互卷积层的输入是两条路径中前一个卷积层输出的串联,交互卷积层通过过滤器同样获得 16 个特征图。两条路径中各自中间两个卷积层的输出参与信息交互卷积层的计算。针对任一条路径,在每个信息交互卷积层计算后,将

所有之前卷积层的输出与此时信息交互卷积层的输出进行串联,其串联结果作为下一卷积层的输入。根据文献[23],为了避免梯度弥散,在任一条路径的前 4 层和两路径之间的交互卷积层均使用批量数据归一化方法,之后再使用 LReLU 激活函数优化网络。生成器网络最后一层的卷积核大小是 1×1 ,连接两条路径中所有卷积层的输出并且将 tanh 作为激活函数。此外,每一层的步长设置为 1,填充模式设置为 SAME。

2.4 鉴别器网络框架

红外鉴别器和可见光鉴别器采用相同的结构。具体地,以红外鉴别器为例,图 4 中鉴别器的输入有两类,一个是红外图像,另一个是融合图像。鉴别器包含前四个卷积层与最后一个线性分类层,4 个卷积层采用 3×3 的卷积核进行卷积计算之后均使用 LReLU 激活函数优化网络,其中后 3 个卷积层也使用了批量数据归一化方法。此外,所有卷积层的填充模式为 VALID,步长设置为 2。最后一个线性层采用 1×1 大小的卷积核,输出红外图像或融合图像的概率。

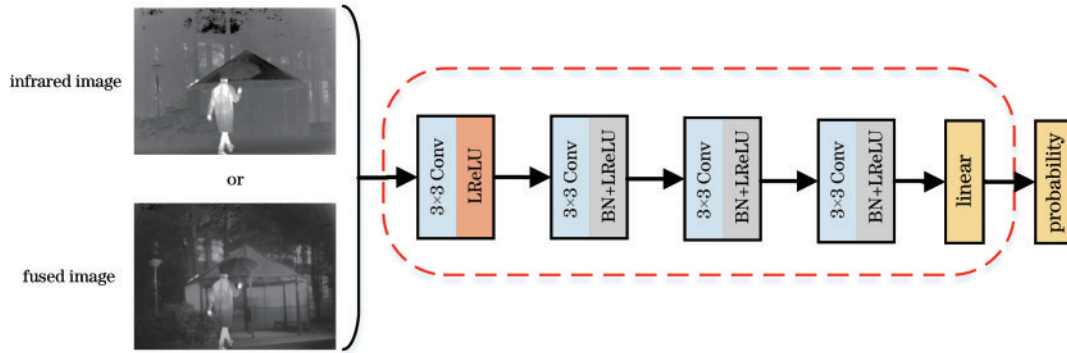


图 4 鉴别器网络框架

Fig. 4 Network architecture of discriminator

2.5 损失函数

损失函数包括生成器损失 L_G 、红外鉴别器损失 L_{Di} 和可见光鉴别器损失 L_{Dv} 。生成器损失包括两部分,其一是在生成器和鉴别器之间的对抗性损失函数 L_{Gadv} ,主要用于约束信息的平衡,另一个是内容损失函数 L_{Gcon} ,主要用于从源图像中提取和重建信息。生成器损失可表示为

$$L_G = L_{Gadv} + \alpha L_{Gcon}, \quad (1)$$

式中, α 是平衡两个损失项的正则化参数。对抗性损失用来指导生成器产生真实的融合图像以欺骗红外鉴别器与可见光鉴别器,内容损失是为了提取并保留源图像中各种有意义的信息,它包含表征热辐射信息的强度损失、自适应增强块在权重图作用下的特征增强损失和表征纹理细节信息的梯度损失,它们的定义如下:

$$L_{Gadv} = \frac{1}{N} \sum_{n=1}^N [P_{Di}(I_f^n) - a_1]^2 + [P_{Dv}(I_f^n) - a_2]^2, \quad (2)$$

$$L_{Gcon} = L_{int} + \beta_1 L_{Ie} + \beta_2 L_{grad}, \quad (3)$$

式中: I_f 是融合图像; N 代表在一个批次的图像训练中融合图像的数量; a_1 与 a_2 分别是红外鉴别器与可见光鉴别器判断融合图像的概率标签; P_{Di} 代表融合图像是红外图像的概率; P_{Dv} 代表融合图像是可见光图像的概率; $\beta(\cdot)$ 是用于调整 3 个损失项使其达到相同重要性级别的正则化参数。由于生成器期望红外鉴别器与可见光鉴别器分别将融合图像误识别成红外图像与可见光图像,因此 a_1 与 a_2 被设置为 1。

强度损失 L_{int} 用于约束融合图像以保持其与源图像具有相同的强度分布,从而保留明显的对比度信息,突出目标区域,其表达式为

$$L_{int} = \frac{1}{HW} \sum_{x=1}^H \sum_{y=1}^W (I_{I,x,y} - I_{V,x,y})^2 + \gamma (I_{I,x,y} - I_{V,x,y})^2, \quad (4)$$

式中: H 和 W 分别表示图像的高和宽; x 和 y 代表图像中第 x 行第 y 列的像素; I_I 和 I_V 分别表示红外与可见光图像; γ 是用于权衡两个损失项的正则化参数。

特征增强损失 L_{fe} 可以使融合图像中来自红外与可见光图像的纹理细节与热目标的边缘得到锐化与增强,其表达式为

$$L_{fe} = \frac{1}{HW} \sum_{x=1}^H \sum_{y=1}^W (I_{I_{x,y}} - I_{V_{x,y}})^2 \cdot \omega_i(x, y) + \gamma (I_{I_{x,y}} - I_{V_{x,y}})^2 \cdot \omega_v(x, y), \quad (5)$$

式中, $\omega(x, y)$ 是在自适应增强块作用下的权重图,其生成过程可以表示为

$$\omega(x, y) = G_{r_1 \times r_1}(\nabla^2 I) + G_{r_2 \times r_2}(\nabla^2 I) + G_{r_3 \times r_3}(\nabla^2 I), \quad (6)$$

式中: G 表示高斯卷积核; $r(\bullet) \times r(\bullet)$ 对应于高斯核的大小,其中, $r_1=3$ 、 $r_2=5$ 、 $r_3=7$; $\nabla^2 I$ 表示红外或者可见光图像经过拉普拉斯梯度算子所得梯度的平方。由于权重图的感兴趣区域在于源图像中梯度值较大的地方,因此在权重图中针对源图像大多数无明显梯度变化的区域权重为零,而在各类物体边缘区域权重很大。此外,高斯核的正态分布特性可以帮助权重值的分布得到平滑过渡。

梯度损失 L_{grad} 用于约束融合图像使其保留源图像中丰富的纹理细节信息,其表达式为

$$L_{grad} = \frac{1}{HW} \sum_{x=1}^H \sum_{y=1}^W (\nabla I_{I_{x,y}} - \nabla I_{V_{x,y}})^2 + \delta (\nabla I_{I_{x,y}} - \nabla I_{I_{x,y}})^2, \quad (7)$$

式中: $\nabla I_{V_{x,y}}$ 、 $\nabla I_{I_{x,y}}$ 和 $\nabla I_{I_{x,y}}$ 分别表示可见光图像的梯度、红外图像的梯度和融合图像的梯度; δ 用来控制两个损失项之间的平衡。

红外鉴别器与可见光鉴别器的损失函数表示为

$$L_{Di} = \frac{1}{N} \sum_{n=1}^N [P_{Di}(I_i^n) - b_1]^2 + [P_{Di}(I_i^n) - b_2]^2, \quad (8)$$

$$L_{Dv} = \frac{1}{N} \sum_{n=1}^N [P_{Dv}(I_v^n) - c_1]^2 + [P_{Dv}(I_i^n) - c_2]^2, \quad (9)$$

式中, b_1 和 b_2 是红外鉴别器分别判断红外图像与融合图像的概率标签,由于红外鉴别器需要正确地判断出红外图像与融合图像,因此 b_1 被设置为 1, b_2 被设置为 0。同理,由于可见光鉴别器也需要正确地判断出可见光图像与融合图像,因此 c_1 被设置为 1, c_2 被设置为 0。

3 实验分析与讨论

3.1 数据集与实验细节

实验是在公开的 TNO 数据集^[24]上进行的, TNO 数据集主要包括军事等相关场所。首先,从 TNO 数据集中选择 40 个已注册的红外与可见光图像对。为了扩充训练数据,通过步长为 14 的滑窗,将输入图像裁剪成 37357 个尺寸为 120×120 的图像块。最后,从 TNO 数据集剩余的图像中选择 18 对红外与可见光图像进行测试。

具体地,鉴别器和生成器根据对抗性原理进行迭代训练,两者的训练次数的比值为 p , 批量大小被设为

b , 完成一次迭代需要 m 步, 训练阶段的总迭代次数设置为 M 。在实验中, 经验地设置 $b=32$ 、 $p=2$ 、 $M=10$, m 是所有图像块数量与 b 的比值, 训练中的参数由 Adam 优化器更新, 整个训练过程如表 1 所示。根据实验中多次调参的结果, 损失项参数的最终结果如下: 式(1)中 $\alpha=100$, 式(3)中 $\beta_1=5$ 、 $\beta_2=7$, 式(4)中 $\gamma=0.3$, 式(7)中 $\delta=0.8$ 。此外, 将应该设置为 1 的概率标签 a_1 、 a_2 、 b_1 和 c_1 放宽至 0.7~1.2 的范围, 将应该设置为 0 的概率标签 b_2 和 c_2 放宽到 0~0.3 的范围。最后, 基于深度学习的训练由 GPU NVIDIA Tesla P100 完成。此外, 为了后续在同等环境下对传统方法与基于深度学习的方法进行比较, 所有的测试实验由 CPU i7-8550U 完成。

表 1 所提方法的训练过程

Table 1 Training process of proposed method

Algorithm: training process of proposed method	
Input:	training set infrared image I_i and visible image I_v
Output:	fused image I_f
1)	for M epochs do
2)	for m steps do
3)	for p times do
4)	Select b infrared patches $\{I_i^1, I_i^2, \dots, I_i^b\}$
5)	Select b visible patches $\{I_v^1, I_v^2, \dots, I_v^b\}$
6)	Select b fused patches $\{I_f^1, I_f^2, \dots, I_f^b\}$
7)	Update the parameters of the discriminator by Adam optimizer
8)	end
9)	Select b infrared patches $\{I_i^1, I_i^2, \dots, I_i^b\}$
10)	Select b visible patches $\{I_v^1, I_v^2, \dots, I_v^b\}$
11)	Update the parameters of the generator by Adam optimizer
12)	end
13)	end

3.2 客观评价指标

选用 6 个指标来客观评价融合图像的性能, 包括熵(EN)^[25]、视觉保真度(VIF)^[26]、相关系数(CC)^[27]、差异相关性之和(SCD)^[28]、多尺度结构相似性(MS_SSIM)^[29]和像素特征互信息(FMI_{pixel})^[30]。

EN 表示融合图像中包含的信息量, EN 的值越大, 表示融合图像包含的信息越多。VIF 衡量融合图像的信息保真度, VIF 的值越大表明融合图像越符合人眼的视觉感知。CC 用来度量融合图像与源图像之间的线性相关性, CC 的值越大, 则表明融合图像与源图像的相关性越强。SCD 用于度量源图像与融合图像之间的相关程度, 并且同样适用于评估伪信息, SCD 的值越大, 表明融合图像与源图像相关程度越高, 伪信息越少。MS_SSIM 是一种多尺度结构相似性度量方法, 比单尺度方法具有更大的灵活性, 可以度量融合图像与源图像之间的结构相似度, 其值越大表明结构保持得越好。FMI_{pixel} 用来评估像素特征互信

息,其值越大表明融合图像包含更多的源图像信息。

3.3 对比实验

将所提方法与 9 种典型的图像融合方法进行了比较,包括 MDLatLRR^[31]、FPDE^[32]、LatLRR^[33]、GTF^[34]、DDcGAN^[19]、DenseFuse^[16]、U2Fusion^[13]、FusionGAN^[18]和 GANMcC^[20]。其中,前 4 种方法是传统的融合方法,后 5 种方法是基于深度学习的融合方法。首先在 bench 图像中进行可视化比较,结果如图 5 所示,实线框标记对比度信息,虚线框标记纹理细节。MDLatLRR、FPDE、LatLRR 和 U2Fusion 的融合图像虽然保留了足够的纹理细节,但是在对比度方面存在着明显的不足。GTF 的融合图像虽然保留了

红外图像的高对比度,但是在可见光图像的纹理细节方面保留不足。DDcGAN 和 FusionGAN 的融合图像虽然对于热辐射目标保持高亮的对比度,但是它们的目标边缘非常模糊。GANMcC 相较于 FusionGAN 而言,其融合图像在一定程度上提升了纹理细节却降低了对比度。DenseFuse 的融合图像偏向于保留可见光图像的纹理细节,虽然有一定的对比度信息,但是边缘轮廓依旧不够清晰,而且来自可见光图像中的木桩轮廓很难辨识。相比之下,所提方法在充分保留纹理细节的同时锐化且增强了来自红外图像的热辐射目标,使得人体的轮廓清晰可见,并且整体高亮突出,效果最佳。

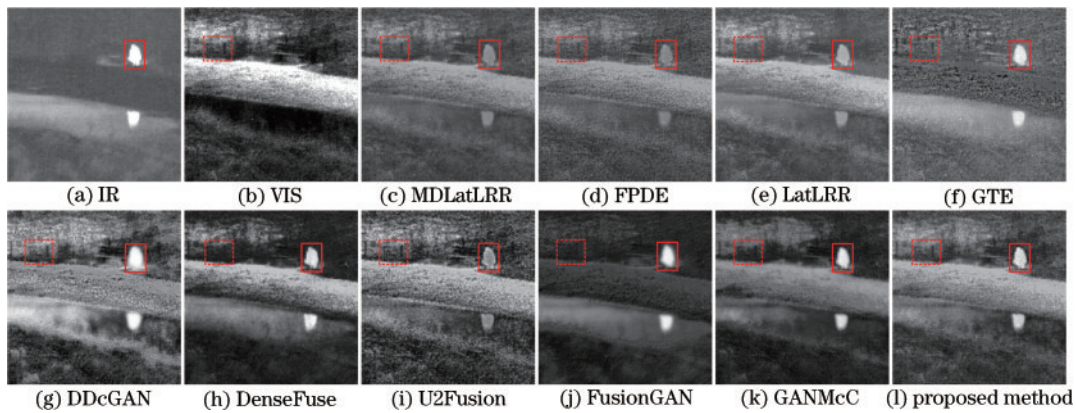


图 5 不同方法在 bench 图像的融合结果

Fig. 5 Fusion results of different methods in bench images

为了进一步探究所提方法的效果,选择另外 6 对图像进行对比,结果如图 6 所示,前两行是初始的红外与可见光图像,所提方法的融合结果显示在最后一行,剩下的对比实验的融合结果分别在第 3~11 行。从融合结果可以看出,与其他方法相比,所提方法在充分保留纹理细节的同时可以锐化增强各类目标,包括红外图像中的热目标与可见光图像中的高对比度目标,从而帮助融合图像实现显著性检测与目标定位的目的。

最后从客观角度分析 TNO 数据集的 18 对测试图像,6 个指标的定量比较结果如图 7 所示,所提方法在 VIF、CC 和 MS_SSIM 上取得了最大平均值。对于 SCD,所提方法仅略微次于 U2Fusion。以上指标表明所提方法更适合人眼的直观视觉感知,融合图像与源图像的线性相关性更强,并且保持良好的结构信息。所提方法的 SCD 数值表明其伪信息相对较少,与 U2Fusion 相比,两者的指标平均值差距很小,但结合 CC,U2Fusion 的线性相关性不足,结合 VIF,所提方法比 U2Fusion 具有更好的视觉保真度。虽然从总体上看,所提方法在 6 个指标上的平均值与 U2Fusion 相比差异不明显,但是结合图 6,针对融合图像中热目标的对比度信息,所提方法明显优于 U2Fusion,而保持高对比度信息正是红外与可见光图像融合的主要目的。综上所述,所提方法优于 U2Fusion。对于 EN,所提方

法在平均值上次于 DDcGAN、DenseFuse 和 U2Fusion;对于 FMI_{pixel} ,仅次于 GTF 和 MDLatLRR。这是因为增强块在训练过程中平滑了源图像各类目标物体边缘的少量稀疏噪点,导致融合图像在 EN 和 FMI_{pixel} 上的平均值相比其他个别方法而言优势不明显。但是所提方法在两个源图像信息之间实现了更好的平衡,融合结果仍然包含了大量的源图像信息,总体而言,所提方法在客观评价中的表现最佳。

3.4 消融实验

自适应增强块分析:特征增强损失函数被引入以加强来自源图像的各类目标物体的边缘与纹理细节。针对式(6),权重图的大部分区域权重为 0,因为这些区域可以在强度损失中被很好地优化。针对源图像中各类目标物体的边缘,权重图对应的权重值较大,远离目标边缘权重值逐渐减小并且平滑过渡为 0,同时,源图像中各类目标物体边缘的杂乱噪点可以得到平滑滤波。为了证明其有效性,设计对应的消融实验。以 Kaptein_1123 图像对为例,去除特征增强损失函数前后分别训练网络,结果如图 8 所示,实线用于标记天空中的噪点,虚线用于标记地面的纹理细节。从融合图像可以看到,特征增强损失函数有效地锐化了源图像中各类目标物体的边缘,在保持可见光图像纹理细节清晰的同时还增强了红外热目标的高对比度。具体



图6 不同方法在6对典型的红外与可见光图像上的定性比较

Fig. 6 Qualitative comparison of different methods on 6 pairs of typical infrared and visible images

地,取消特征增强损失函数之后,在 Kaptein_1123 融合图像的天空中有来自红外图像的杂乱且分布不均的噪点,而地砖的条纹也变得离散且杂乱。这表明特征增强损失函数很好地改善了融合图像的质量,加强了融合图像中来自源图像的纹理细节和热目标的边缘。

为了进一步说明自适应增强块中权重图的作用,图9详细地展示了权重作用前后的效果差异,其原理在于式(6)的权重图采用不同半径组合的高斯卷积核对源图像的梯度图进行滤波。具体地,半径分别为3、5和7的高斯卷积核被组合,随后对源图像的梯度图进行高斯滤波。在图9中:针对红外图像,与其梯度图相

比,权重作用后的锐化效果图中热目标的边缘纹理得到了明显的增强,例如实线框中行人的轮廓变得更加清晰和高亮。针对可见光图像,在权重作用后的锐化效果图中,纹理细节十分突出,例如虚线框中的地砖条纹、帐篷支架和树枝等纹理细节得到了锐化,同时,其他杂乱且离散的噪点可以被平滑滤波。

超参数分析:所提模型中有 α 、 β_1 、 β_2 、 γ 和 δ 等5个超参数。其中, α 用来权衡内容损失和对抗性损失,然而对抗性机制中 α 只需要保证鉴别器进行有效对抗即可,其值变化不会对模型产生太大的影响。 γ 用来权衡强度损失中红外与可见光图像各自的对比度信息,

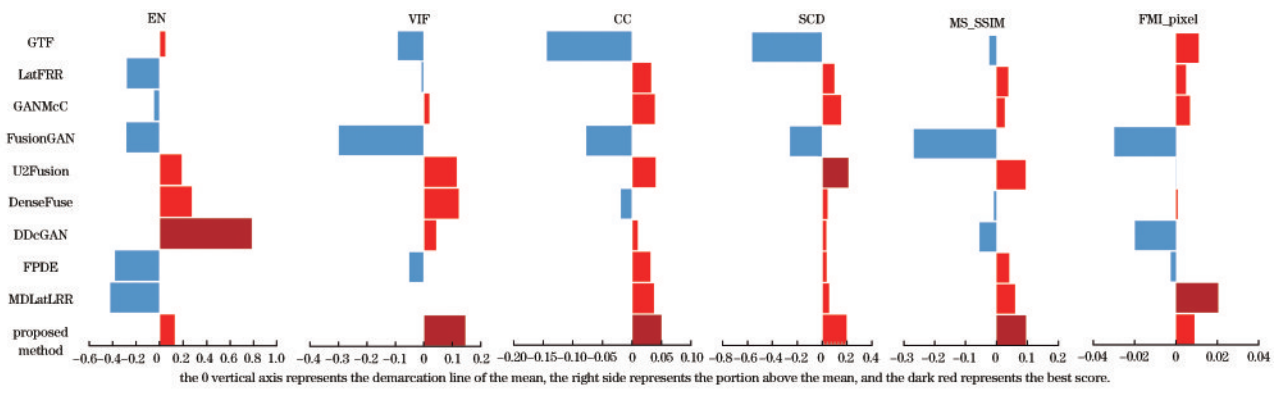


图 7 6 个指标在 18 对 TNO 测试集的定量比较

Fig. 7 Quantitative comparison of six indicators in 18 pairs of TNO test sets

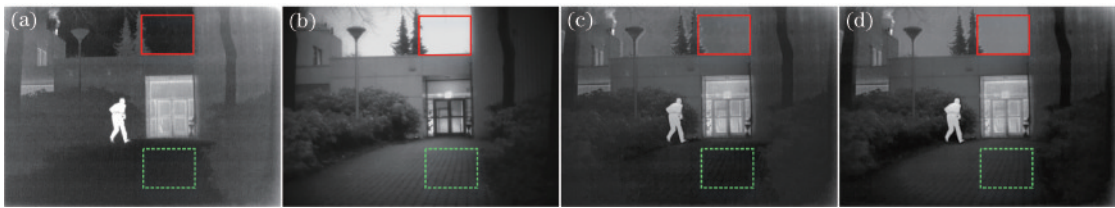


图 8 Kaptein_1123 图像针对增强块的消融实验。(a) 红外图像; (b) 可见光图像; (c) 去除增强块的融合图像; (d) 所提方法的融合图像
Fig. 8 Kaptein_1123 images for ablation experiments of enhanced blocks. (a) Infrared image; (b) visible image; (c) fused image without the enhancement block; (d) fused image of the proposed method

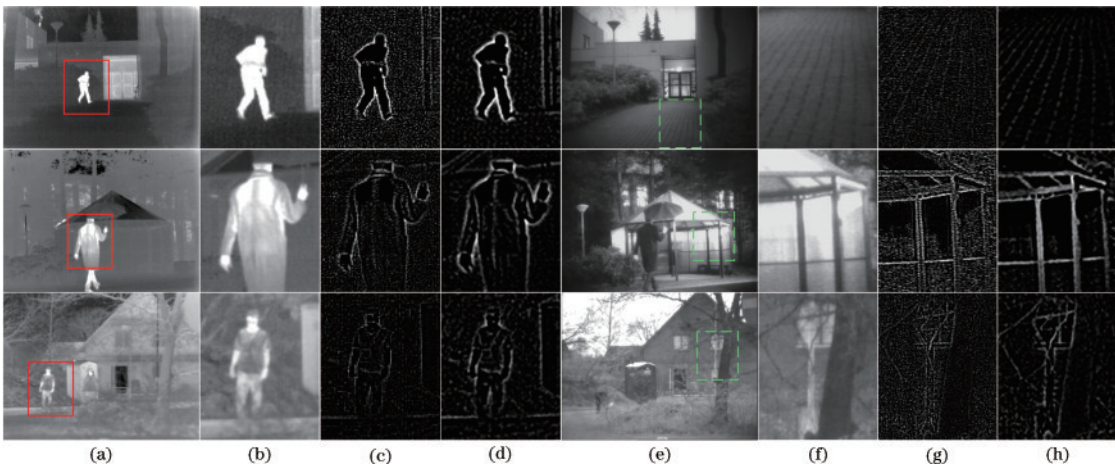


图 9 权重作用前后的效果。(a) 红外图像; (b) 局部红外图像; (c) 局部红外梯度图; (d) 权重作用于红外梯度的效果; (e) 可见光图像; (f) 局部可见光图像; (g) 局部可见光梯度图; (h) 权重作用于可见光梯度的效果
Fig. 9 Effects before and after weight action. (a) Infrared image; (b) local infrared image; (c) local infrared gradient map; (d) effect of weights acting on infrared gradient; (e) visible image; (f) local visible image; (g) local visible gradient map; (h) effect of weights acting on visible gradient

δ 用来权衡梯度损失中红外与可见光图像各自的纹理细节, 由于对比度主要来自于红外图像, 纹理细节主要来自可见光图像, 因此经验地设置 γ 与 δ 分别为 0.3 和 0.8 即可。综上所述, 只需对 β_1 和 β_2 进行敏感性分析。

根据文献[20], 梯度损失项系数为强度损失项系数的 5 倍, 而在本研究中特征增强损失是由自适应增强块的权重图和强度损失联合构成的, 这间接地导致式(3)中强度损失的比重有所增加。为了平衡内容损

失函数的 3 个损失项, 需要在文献[20]的参数设置的基础上适当加大梯度损失项系数 β_2 。因此, 首先预设 β_2 的值为 7 并且固定 β_2 , 网络采用不同的特征增强损失项系数 β_1 的值分别进行训练, 结果如图 10 第 1 行和第 3 行所示, 从左至右, β_1 的值逐步递增。具体地, 当 $\beta_1=1$ 时自适应增强块的作用很微弱, 融合结果中纹理细节和热目标的边缘较为模糊。在 β_1 递增的过程中, 融合结果的纹理细节和热目标的边缘逐渐清晰且呈现锐化效果。当 $\beta_1=5$ 时, 融合结果达到了最佳状态, 在

保持锐化效果的同时,其强度分布和梯度分布达到了良好的平衡;当 β_1 大于7时,例如在 $\beta_1=9$ 的融合结果中,虽然物体的边缘纹理得到了锐化与增强,但是由于权重图作用于强度损失,间接地导致强度损失作用增强而梯度损失作用减弱,融合结果偏向源图像的强度分布。因此 β_1 的最佳取值固定为5,随即网络围绕 β_2 的不同取值进行训练,结果如图10第2行和第4行所示,从左至右, β_2 的值逐步递增。具体地,当 $\beta_2=2$ 时,融合结果偏向源图像的强度分布,在视

觉效果上类似于红外图像。在 β_2 的递增过程中,融合结果中热目标的对比度逐渐下降,如红色实线框中的行人和地堡。当 $\beta_2=7$ 时,融合结果良好地控制了强度分布和梯度分布的平衡,热目标和纹理细节均突出;当 β_2 大于10时,融合结果中热目标的对比度呈现快速下降趋势;当 $\beta_2=13$ 时,融合结果中热目标的对比度明显减弱,在视觉效果上偏向可见光图像。综上所述, β_1 和 β_2 分别被设置为5和7时,融合图像的效果最佳。

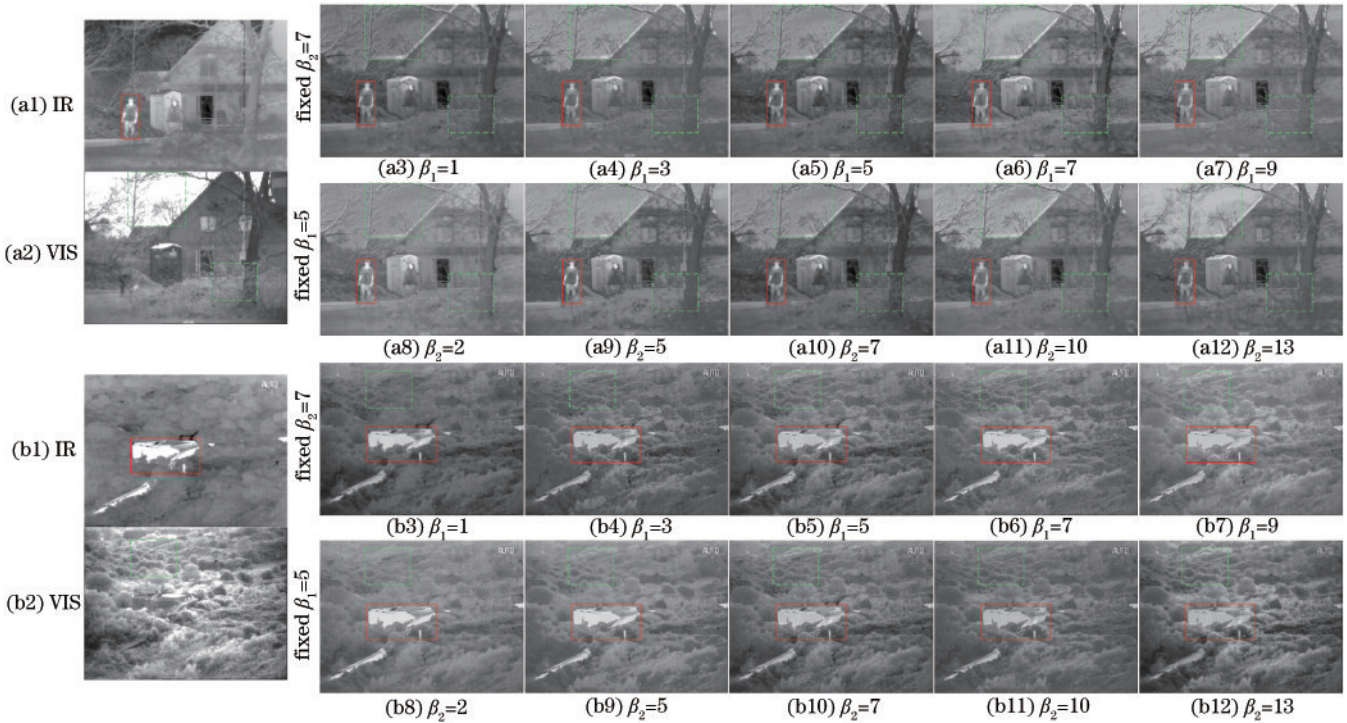


图 10 超参数的消融实验分析

Fig. 10 Experimental analysis of ablation with hyperparameters

3.5 运行效率

由于生成器是所提方法的目标网络,为了进一步验证目标网络的效率,表2提供了不同方法在每张测试图像上的平均运行时间。测试时,每种方法的输入输出均一致。同时为了使得测试结果更客观,所有方法在运行时统一使用1.80 GHz Intel Core i7-8550U CPU,以保证硬件条件相同。可以看出,所提方法在运行效率上是很可观的,与其他方法相比速度最快。

表 2 不同方法在所选的 18 对 TNO 测试集上的平均运行时间
Table 2 Average running time of different methods on the selected 18 pairs of TNO test set unit: s

Method	TNO	Method	TNO
MDLatLRR	70.5981	DenseFuse	5.0580
FPDE	2.9546	U2Fusion	6.0424
LatLRR	154.4078	FusionGAN	7.1445
GTF	6.8297	GANMcC	14.3756
DDcGAN	15.6905	Proposed method	1.6773

4 结 论

提出一种基于自适应特征增强和生成器路径交互的生成对抗网络,用于红外与可见光图像融合。该方法是一个端到端的模型,由高斯卷积核过滤源图像得到权重图的自适应增强块与强度损失进行联合约束,可以动态地引导网络产生的融合图像锐化并增强纹理细节和热目标的边缘,从而使得融合结果的边缘纹理清晰可见。此外,生成器的内部结构也进行了相应的调整和改善,在其两条路径之间进行信息交互,这样可以更充分地提取源图像的各类信息。以主辅思想设计的损失函数和网络结构中的双鉴别器进一步帮助融合图像平衡对比度与纹理细节。所提方法与其他9种典型的图像融合方法的定性和定量实验结果表明,所提方法均优于其他方法,其融合结果具有轮廓清晰的热目标与纹理细节信息。此外,所提方法还在运行效率上面表现最佳、速度最快。

参 考 文 献

- [1] Piella G. A general framework for multiresolution image fusion: from pixels to regions[J]. *Information Fusion*, 2003, 4(4): 259-280.
- [2] 张寅, 朱桂熠, 施天俊, 等. 基于特征融合与注意力的遥感图像小目标检测[J]. *光学学报*, 2022, 42(24): 2415001.
Zhang Y, Zhu G Y, Shi T J, et al. Small object detection in remote sensing images based on feature fusion and attention[J]. *Acta Optica Sinica*, 2022, 42(24): 2415001.
- [3] 何自芬, 陈光晨, 陈俊松, 等. 多尺度特征融合轻量化夜间红外行人实时检测[J]. *中国激光*, 2022, 49(17): 1709002.
He Z F, Chen G C, Chen J S, et al. Multi-scale feature fusion lightweight real-time infrared pedestrian detection at night[J]. *Chinese Journal of Lasers*, 2022, 49(17): 1709002.
- [4] 王乐, 方玥, 王胜春, 等. 基于偏振融合的钢轨廓形线结构光成像方法[J]. *光学学报*, 2020, 40(22): 2211001.
Wang L, Fang Y, Wang S C, et al. Line-structured light imaging method of rail profile based on polarization fusion[J]. *Acta Optica Sinica*, 2020, 40(22): 2211001.
- [5] Li S T, Kang X D, Fang L Y, et al. Pixel-level image fusion: a survey of the state of the art[J]. *Information Fusion*, 2017, 33: 100-112.
- [6] 肖万新, 李华锋, 张亚飞, 等. 多尺度特征学习和边缘增强的医学图像融合[J]. *激光与光电子学进展*, 2022, 59(6): 0617029.
Xiao W X, Li H F, Zhang Y F, et al. Medical image fusion based on multi-scale feature learning and edge enhancement[J]. *Laser & Optoelectronics Progress*, 2022, 59(6): 0617029.
- [7] Chen J, Li X J, Luo L B, et al. Infrared and visible image fusion based on target-enhanced multiscale transform decomposition[J]. *Information Sciences*, 2020, 508: 64-78.
- [8] Yang Y, Zhang Y M, Huang S Y, et al. Infrared and visible image fusion using visual saliency sparse representation and detail injection model[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 5001715.
- [9] 裴佩佩, 杨艳春, 党建武, 等. 基于滚动引导滤波器和卷积稀疏表示的红外与可见光图像融合方法[J]. *激光与光电子学进展*, 2022, 59(12): 1210001.
Pei P P, Yang Y C, Dang J W, et al. Infrared and visible image fusion method based on rolling guidance filter and convolution sparse representation[J]. *Laser & Optoelectronics Progress*, 2022, 59(12): 1210001.
- [10] 张立保, 章珏. 基于显著性分析的自适应遥感图像融合[J]. *中国激光*, 2015, 42(1): 0114001.
Zhang L B, Zhang J. A new adaptive fusion method based on saliency analysis for remote sensing images[J]. *Chinese Journal of Lasers*, 2015, 42(1): 0114001.
- [11] Wang Z S, Xu J W, Jiang X L, et al. Infrared and visible image fusion via hybrid decomposition of NSCT and morphological sequential toggle operator[J]. *Optik*, 2020, 201: 163497.
- [12] Paramanandham N, Rajendiran K. Multi sensor image fusion for surveillance applications using hybrid image fusion algorithm[J]. *Multimedia Tools and Applications*, 2018, 77(10): 12405-12436.
- [13] Xu H, Ma J Y, Jiang J J, et al. U2Fusion: a unified unsupervised image fusion network[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(1): 502-518.
- [14] Zhang H, Ma J Y. SDNet: a versatile squeeze-and-decomposition network for real-time image fusion[J]. *International Journal of Computer Vision*, 2021, 129(10): 2761-2785.
- [15] 陈国洋, 吴小俊, 徐天阳. 基于深度学习的无监督红外图像与可见光图像融合算法[J]. *激光与光电子学进展*, 2022, 59(4): 0410010.
Chen G Y, Wu X J, Xu T Y. Unsupervised infrared image and visible image fusion algorithm based on deep learning[J]. *Laser & Optoelectronics Progress*, 2022, 59(4): 0410010.
- [16] Li H, Wu X J. DenseFuse: a fusion approach to infrared and visible images[J]. *IEEE Transactions on Image Processing*, 2019, 28(5): 2614-2623.
- [17] Zhang Y, Liu Y, Sun P, et al. IFCNN: a general image fusion framework based on convolutional neural network[J]. *Information Fusion*, 2020, 54: 99-118.
- [18] Ma J Y, Yu W, Liang P W, et al. FusionGAN: a generative adversarial network for infrared and visible image fusion[J]. *Information Fusion*, 2019, 48: 11-26.
- [19] Ma J, Xu H, Jiang J, et al. DDcGAN: a dual-discriminator conditional generative adversarial network for multi-resolution image fusion[J]. *IEEE Transactions on Image Processing*, 2020, 29: 4980-4995.
- [20] Ma J Y, Zhang H, Shao Z F, et al. GANMcC: a generative adversarial network with multiclassification constraints for infrared and visible image fusion[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 1-14.
- [21] Zhang H, Le Z L, Shao Z F, et al. MFF-GAN: an unsupervised generative adversarial network with adaptive and gradient joint constraints for multi-focus image fusion[J]. *Information Fusion*, 2021, 66: 40-53.
- [22] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2261-2269.
- [23] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[EB/OL]. (2015-11-19) [2022-08-02]. <https://arxiv.org/abs/1511.06434>.
- [24] Toet A. TNO image fusion dataset[DB/OL]. (2014-04-26) [2022-08-02]. https://figshare.com/articles/dataset/TNO_Image_Fusion_Dataset/1008029.
- [25] Roberts J W, van Aardt J A, Ahmed F B. Assessment of image fusion procedures using entropy, image quality,

- and multispectral classification[J]. *Journal of Applied Remote Sensing*, 2008, 2(1): 023522.
- [26] Han Y, Cai Y Z, Cao Y, et al. A new image fusion performance metric based on visual information fidelity [J]. *Information Fusion*, 2013, 14(2): 127-135.
- [27] Deshmukh M, Bhosale U. Image fusion and image quality assessment of fused images[J]. *International Journal of Image Processing (IJIP)*, 2010, 4(5): 484-508.
- [28] Aslantas V, Bendes E. A new image quality metric for image fusion: the sum of the correlations of differences [J]. *AEU-International Journal of Electronics and Communications*, 2015, 69(12): 1890-1896.
- [29] Wang Z, Simoncelli E P, Bovik A C. Multiscale structural similarity for image quality assessment[C]// *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers*, November 9-12, 2003, Pacific Grove, CA, USA. New York: IEEE Press, 2003: 1398-1402.
- [30] Haghghat M, Razian M A. Fast-FMI: Non-reference image fusion metric[C]//2014 IEEE 8th International Conference on Application of Information and Communication Technologies, October 15-17, 2014, Astana, Kazakhstan. New York: IEEE Press, 2014.
- [31] Li H, Wu X J, Kittler J. MDLatLRR: a novel decomposition method for infrared and visible image fusion[J]. *IEEE Transactions on Image Processing*, 2020, 29: 4733-4746.
- [32] Bavirisetti D P, Xiao G, Liu G. Multi-sensor image fusion based on fourth order partial differential equations [C]//2017 20th International Conference on Information Fusion (Fusion), July 10-13, 2017, Xi'an, China. New York: IEEE Press, 2017.
- [33] Li H, Wu X J. Infrared and visible image fusion using latent low-rank representation[EB/OL]. (2018-04-24) [2022-08-02]. <https://arxiv.org/abs/1804.08992>.
- [34] Ma J Y, Chen C, Li C, et al. Infrared and visible image fusion via gradient transfer and total variation minimization[J]. *Information Fusion*, 2016, 31: 100-109.